(1) GENERAL INFORMATION:

    (i) APPLICANT: ANDERSON, Darrell R.
                 HANNA, Nabil
                 LEONARD, John E.
                 NEWMAN, Roland A.
                 REFF, Mitchell E.
                 RASTETTER, William H.

    (ii) TITLE OF INVENTION: THERAPEUTIC APPLICATION OF CHIMERIC AND
        RADIOLABELED ANTIBODIES TO HUMAN B LYMPHOCYTE RESTRICTED
        DIFFERENTIATION ANTIGEN FOR THE TREATMENT OF B CELL
        LYMPHOMA

    (iii) NUMBER OF SEQUENCES: 9

    (iv) CORRESPONDENCE ADDRESS:
        (A) ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
        (B) STREET: P.O. Box 1404
        (C) CITY: Alexandria
        (D) STATE: Virginia
        (E) COUNTRY: United States
        (F) ZIP: 22313-1404

    (v) COMPUTER READABLE FORM:
        (A) MEDIUM TYPE: Floppy disk
        (B) COMPUTER: IBM PC compatible
        (C) OPERATING SYSTEM: PC-DOS/MS-DOS
        (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

    (vi) CURRENT APPLICATION DATA:
        (A) APPLICATION NUMBER: US 08/149,099
        (B) FILING DATE: 03-NOV-1993
        (C) CLASSIFICATION:

    (vii) PRIOR APPLICATION DATA:
        (A) APPLICATION NUMBER: US 07/978,891
        (B) FILING DATE: 12-NOV-1992

    (viii) ATTORNEY/AGENT INFORMATION:
        (A) NAME: Teskin, Robin L.
        (B) REGISTRATION NUMBER: 35,030
        (C) REFERENCE/DOCKET NUMBER: 012712-014

    (ix) TELECOMMUNICATION INFORMATION:
        (A) TELEPHONE: (703) 836-6620
        (B) TELEFAX: (703) 836-2021

(2) INFORMATION FOR SEQ ID NO:1:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 27 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GGGAGCTTGG ATCGATCCTC TATGGTT

27

(2) INFORMATION FOR SEQ ID NO:2:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 8540 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: circular

    (ii) MOLECULE TYPE: DNA (genomic)

    (iii) HYPOTHETICAL: NO

    (iv) ANTI-SENSE: NO

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
GACGTCGCGG CCGCTCTAGG CCTCCAAAAA AGCCTCCTCA CTACTTCTGG AATAGCTCAG      60

AGGCCGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAAT TAGTCAGCCA TGCATGGGGC     120

GGAGAATGGG CGGAACTGGG CGGAGTTAGG GGCGGGATGG GCGGAGTTAG GGGCGGGACT     180

ATGGTTGCTG ACTAATTGAG ATGCATGCTT TGCATACTTC TGCCTGCTGG GGAGCCTGGG     240

GACTTTCCAC ACCTGGTTGC TGACTAATTG AGATGCATGC TTTGCATACT TCTGCCTGCT     300

GGGGAGCCTG GGGACTTTCC ACACCCTAAC TGACACACAT TCCACAGAAT TAATTCCCCT     360

AGTTATTAAT AGTAATCAAT TACGGGGTCA TTAGTTCATA GCCCATATAT GGAGTTCCGC     420

GTTACATAAC TTACGGTAAA TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCCATTG     480

ACGTCAATAA TGACGTATGT TCCCATAGTA ACGCCAATAG GGACTTTCCA TTGACGTCAA     540

TGGGTGGACT ATTTACGGTA AACTGCCCAC TTGGCAGTAC ATCAAGTGTA TCATATGCCA     600

AGTACGCCCC CTATTGACGT CAATGACGGT AAATGGCCCG CCTGGCATTA TGCCCAGTAC     660

ATGACCTTAT GGGACTTTCC TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC     720

ATGGTGATGC GGTTTTGGCA GTACATCAAT GGGCGTGGAT AGCGGTTTGA CTCACGGGGA     780

TTTCCAAGTC TCCACCCCAT TGACGTCAAT GGGAGTTTGT TTTGGCACCA AAATCAACGG     840

GACTTTCCAA AATGTCGTAA CAACTCCGCC CCATTGACGC AAATGGGCGG TAGGCGTGTA     900

CGGTGGGAGG TCTATATAAG CAGAGCTGGG TACGTGAACC GTCAGATCGC CTGGAGACGC     960

CATCACAGAT CTCTCACCAT GAGGGTCCCC GCTCAGCTCC TGGGGCTCCT GCTGCTCTGG    1020

CTCCCAGGTG CACGATGTGA TGGTACCAAG GTGGAAATCA AACGTACGGT GGCTGCACCA    1080

TCTGTCTTCA TCTTCCCGCC ATCTGATGAG CAGTTGAAAT CTGGAACTGC CTCTGTTGTG    1140

TGCCTGCTGA ATAACTTCTA TCCCAGAGAG GCCAAAGTAC AGTGGAAGGT GGATAACGCC    1200

CTCCAATCGG GTAACTCCCA GGAGAGTGTC ACAGAGCAGG ACAGCAAGGA CAGCACCTAC    1260

AGCCTCAGCA GCACCCTGAC GCTGAGCAAA GCAGACTACG AGAAACACAA AGTCTACGCC    1320

TGCGAAGTCA CCCATCAGGG CCTGAGCTCG CCCGTCACAA AGAGCTTCAA CAGGGGAGAG    1380

TGTTGAATTC AGATCCGTTA ACGGTTACCA ACTACCTAGA CTGGATTCGT GACAACATGC    1440

GGCCGTGATA TCTACGTATG ATCAGCCTCG ACTGTGCCTT CTAGTTGCCA GCCATCTGTT    1500

GTTTGCCCCT CCCCCGTGCC TTCCTTGACC CTGGAAGGTG CCACTCCCAC TGTCCTTTCC    1560
```

- 67 -

```
AGACAGTGTT CTCTGCACAG ATAAGGACAA ACATTATTCA GAGGGAGTAC CCAGAGCTGA   ·3660 .

GACTCCTAAG CCAGTGAGTG GCACAGCATT CTAGGGAGAA ATATGCTTGT CATCACCGAA    3720

GCCTGATTCC GTAGAGCCAC ACCTTGGTAA GGGCCAATCT GCTCACACAG GATAGAGAGG    3780

GCAGGAGCCA GGGCAGAGCA TATAAGGTGA GGTAGGATCA GTTGCTCCTC ACATTTGCTT    3840

CTGACATAGT TGTGTTGGGA GCTTGGATAG CTTGGACAGC TCAGGGCTGC GATTTCGCGC    3900

CAAACTTGAC GGCAATCCTA GCGTGAAGGC TGGTAGGATT TTATCCCCGC TGCCATCATG   ·3960

GTTCGACCAT TGAACTGCAT CGTCGCCGTG TCCCAAAATA TGGGGATTGG CAAGAACGGA    4020

GACCTACCCT GGCCTCCGCT CAGGAACGAG TTCAAGTACT TCCAAAGAAT GACCACAACC    4080

TCTTCAGTGG AAGGTAAACA GAATCTGGTG ATTATGGGTA GGAAAACCTG GTTCTCCATT    4140

CCTGAGAACA ATCGACCTTT AAAGGACAGA ATTAATATAG TTCTCAGTAG AGAACTCAAA    4200

GAACCACCAC GAGGAGCTCA TTTTCTTGCC AAAAGTTTGG ATGATGCCTT AAGACTTATT    4260

GAACAACCGG AATTGGCAAG TAAAGTAGAC ATGGTTTGGA TAGTCGGAGG CAGTTCTGTT    4320

TACCAGGAAG CCATGAATCA ACCAGGCCAC CTTAGACTCT TTGTGACAAG GATCATGCAG    4380

GAATTTGAAA GTGACACGTT TTTCCCAGAA ATTGATTTGG GGAAATATAA ACTTCTCCCA    4440

GAATACCCAG GCGTCCTCTC TGAGGTCCAG GAGGAAAAAG GCATCAAGTA TAAGTTTGAA    4500

GTCTACGAGA AGAAAGACTA ACAGGAAGAT GCTTTCAAGT TCTCTGCTCC CCTCCTAAAG    4560

TCATGCATTT TTATAAGACC ATGGGACTTT TGCTGGCTTT AGATCAGCCT CGACTGTGCC    4620

TTCTAGTTGC CAGCCATCTG TTGTTTGCCC CTCCCCCGTG CCTTCCTTGA CCCTGGAAGG    4680

TGCCACTCCC ACTGTCCTTT CCTAATAAAA TGAGGAAATT GCATCGCATT GTCTGAGTAG    4740

GTGTCATTCT ATTCTGGGGG GTGGGGTGGG GCAGGACAGC AAGGGGGAGG ATTGGGAAGA    4800

CAATAGCAGG CATGCTGGGG ATGCGGTGGG CTCTATGGAA CCAGCTGGGG CTCGAGCTAC    4860

TAGCTTTGCT TCTCAATTTC TTATTTGCAT AATGAGAAAA AAAGGAAAAT TAATTTTAAC    4920

ACCAATTCAG TAGTTGATTG AGCAAATGCG TTGCCAAAAA GGATGCTTTA GAGACAGTGT    4980

TCTCTGCACA GATAAGGACA AACATTATTC AGAGGGAGTA CCCAGAGCTG AGACTCCTAA    5040

GCCAGTGAGT GGCACAGCAT TCTAGGGAGA AATATGCTTG TCATCACCGA AGCCTGATTC    5100

CGTAGAGCCA CACCTTGGTA AGGGCCAATC TGCTCACACA GGATAGAGAG GGCAGGAGCC    5160

AGGGCAGAGC ATATAAGGTG AGGTAGGATC AGTTGCTCCT CACATTTGCT TCTGACATAG    5220

TTGTGTTGGG AGCTTGGATC GATCCTCTAT GGTTGAACAA GATGGATTGC ACGCAGGTTC    5280

TCCGGCCGCT TGGGTGGAGA GGCTATTCGG CTATGACTGG GCACAACAGA CAATCGGCTG    5340

CTCTGATGCC GCCGTGTTCC GGCTGTCAGC GCAGGGGCGC CCGGTTCTTT TTGTCAAGAC    5400

CGACCTGTCC GGTGCCCTGA ATGAACTGCA GGACGAGGCA GCGCGGCTAT CGTGGCTGGC    5460

CACGACGGGC GTTCCTTGCG CAGCTGTGCT CGACGTTGTC ACTGAAGCGG GAAGGGACTG    5520

GCTGCTATTG GGCGAAGTGC CGGGGCAGGA TCTCCTGTCA TCTCACCTTG CTCCTGCCGA    5580

GAAAGTATCC ATCATGGCTG ATGCAATGCG GCGGCTGCAT ACGCTTGATC CGGCTACCTG    5640
```

```
TAATAAAATG AGGAAATTGC ATCGCATTGT CTGAGTAGGT GTCATTCTAT TCTGGGGGGT    ·1620 .

GGGGTGGGGC AGGACAGCAA GGGGGAGGAT TGGGAAGACA ATAGCAGGCA TGCTGGGGAT     1680

GCGGTGGGCT CTATGGAACC AGCTGGGGCT CGACAGCTAT GCCAAGTACG CCCCCTATTG     1740

ACGTCAATGA CGGTAAATGG CCCGCCTGGC ATTATGCCCA GTACATGACC TTATGGGACT     1800

TTCCTACTTG GCAGTACATC TACGTATTAG TCATCGCTAT TACCATGGTG ATGCGGTTTT     1860

GGCAGTACAT CAATGGGCGT GGATAGCGGT TTGACTCACG GGATTTCCA AGTCTCCACC      1920

CCATTGACGT CAATGGGAGT TTGTTTTGGC ACCAAAATCA ACGGGACTTT CCAAAATGTC     1980

GTAACAACTC CGCCCCATTG ACGCAAATGG GCGGTAGGCG TGTACGGTGG GAGGTCTATA     2040

TAAGCAGAGC TGGGTACGTC CTCACATTCA GTGATCAGCA CTGAACACAG ACCCGTCGAC     2100

ATGGGTTGGA GCCTCATCTT GCTCTTCCTT GTCGCTGTTG CTACGCGTGT CGCTAGCACC     2160

AAGGGCCCAT CGGTCTTCCC CCTGGCACCC TCCTCCAAGA GCACCTCTGG GGGCACAGCG     2220

GCCCTGGGCT GCCTGGTCAA GGACTACTTC CCCGAACCGG TGACGGTGTC GTGGAACTCA     2280

GGCGCCCTGA CCAGCGGCGT GCACACCTTC CCGGCTGTCC TACAGTCCTC AGGACTCTAC     2340

TCCCTCAGCA GCGTGGTGAC CGTGCCCTCC AGCAGCTTGG CACCCAGAC CTACATCTGC      2400

AACGTGAATC ACAAGCCCAG CAACACCAAG GTGGACAAGA AAGCAGAGCC CAAATCTTGT     2460

GACAAAACTC ACACATGCCC ACCGTGCCCA GCACCTGAAC TCCTGGGGGG ACCGTCAGTC     2520

TTCCTCTTCC CCCCAAAACC CAAGGACACC CTCATGATCT CCCGGACCCC TGAGGTCACA     2580

TGCGTGGTGG TGGACGTGAG CCACGAAGAC CCTGAGGTCA AGTTCAACTG GTACGTGGAC     2640

GGCGTGGAGG TGCATAATGC CAAGACAAAG CCGCGGGAGG AGCAGTACAA CAGCACGTAC     2700

CGTGTGGTCA GCGTCCTCAC CGTCCTGCAC CAGGACTGGC TGAATGGCAA GGACTACAAG     2760

TGCAAGGTCT CCAACAAAGC CCTCCCAGCC CCCATCGAGA AAACCATCTC CAAAGCCAAA     2820

GGGCAGCCCC GAGAACCACA GGTGTACACC CTGCCCCCAT CCCGGGATGA GCTGACCAGG     2880

AACCAGGTCA GCCTGACCTG CCTGGTCAAA GGCTTCTATC CCAGCGACAT CGCCGTGGAG     2940

TGGGAGAGCA ATGGGCAGCC GGAGAACAAC TACAAGACCA CGCCTCCCGT GCTGGACTCC     3000

GACGGCTCCT TCTTCCTCTA CAGCAAGCTC ACCGTGGACA AGAGCAGGTG GCAGCAGGGG     3060

AACGTCTTCT CATGCTCCGT GATGCATGAG GCTCTGCACA ACCACTACAC GCAGAAGAGC     3120

CTCTCCCTGT CTCCGGGTAA ATGAGGATCC GTTAACGGTT ACCAACTACC TAGACTGGAT     3180

TCGTGACAAC ATGCGGCCGT GATATCTACG TATGATCAGC CTCGACTGTG CCTTCTAGTT     3240

GCCAGCCATC TGTTGTTTGC CCCTCCCCCG TGCCTTCCTT GACCCTGGAA GGTGCCACTC     3300

CCACTGTCCT TTCCTAATAA AATGAGGAAA TTGCATCGCA TTGTCTGAGT AGGTGTCATT     3360

CTATTCTGGG GGGTGGGGTG GGGCAGGACA GCAAGGGGGA GGATTGGGAA GACAATAGCA     3420

GGCATGCTGG GGATGCGGTG GGCTCTATGG AACCAGCTGG GGCTCGACAG CGCTGGATCT     3480

CCCGATCCCC AGCTTTGCTT CTCAATTTCT TATTTGCATA ATGAGAAAAA AAGGAAAATT     3540

AATTTTAACA CCAATTCAGT AGTTGATTGA GCAAATGCGT TGCCAAAAAG GATGCTTTAG     3600
```

```
CCCATTCGAC CACCAAGCGA AACATCGCAT CGAGCGAGCA CGTACTCGGA TGGAAGCCGG    5700
TCTTGTCGAT CAGGATGATC TGGACGAAGA GCATCAGGGG CTCGCGCCAG CCGAACTGTT    5760
CGCCAGGCTC AAGGCGCGCA TGCCCGACGG CGAGGATCTC GTCGTGACCC ATGGCGATGC    5820
CTGCTTGCCG AATATCATGG TGGAAAATGG CCGCTTTTCT GGATTCATCG ACTGTGGCCG    5880
GCTGGGTGTG GCGGACCGCT ATCAGGACAT AGCGTTGGCT ACCCGTGATA TTGCTGAAGA    5940
GCTTGGCGGC GAATGGGCTG ACCGCTTCCT CGTGCTTTAC GGTATCGCCG CTTCCCGATT    6000
CGCAGCGCAT CGCCTTCTAT CGCCTTCTTG ACGAGTTCTT CTGAGCGGGA CTCTGGGGTT    6060
CGAAATGACC GACCAAGCGA CGCCCAACCT GCCATCACGA GATTTCGATT CCACCGCCGC    6120
CTTCTATGAA AGGTTGGGCT TCGGAATCGT TTTCCGGGAC GCCGGCTGGA TGATCCTCCA    6180
GCGCGGGGAT CTCATGCTGG AGTTCTTCGC CCACCCCAAC TTGTTTATTG CAGCTTATAA    6240
TGGTTACAAA TAAAGCAATA GCATCACAAA TTTCACAAAT AAAGCATTTT TTTCACTGCA    6300
TTCTAGTTGT GGTTTGTCCA AACTCATCAA TCTATCTTAT CATGTCTGGA TCGCGGCCGC    6360
GATCCCGTCG AGAGCTTGGC GTAATCATGG TCATAGCTGT TTCCTGTGTG AAATTGTTAT    6420
CCGCTCACAA TTCCACACAA CATACGAGCC GGAGCATAAA GTGTAAAGCC TGGGGTGCCT    6480
AATGAGTGAG CTAACTCACA TTAATTGCGT TGCGCTCACT GCCCGCTTTC CAGTCGGGAA    6540
ACCTGTCGTG CCAGCTGCAT TAATGAATCG GCCAACGCGC GGGGAGAGGC GGTTTGCGTA    6600
TTGGGCGCTC TTCCGCTTCC TCGCTCACTG ACTCGCTGCG CTCGGTCGTT CGGCTGCGGC    6660
GAGCGGTATC AGCTCACTCA AAGGCGGTAA TACGGTTATC CACAGAATCA GGGGATAACG    6720
CAGGAAAGAA CATGTGAGCA AAAGGCCAGC AAAAGGCCAG GAACCGTAAA AAGGCCGCGT    6780
TGCTGGCGTT TTTCCATAGG CTCCGCCCCC CTGACGAGCA TCACAAAAAT CGACGCTCAA    6840
GTCAGAGGTG GCGAAACCCG ACAGGACTAT AAAGATACCA GGCGTTTCCC CCTGGAAGCT    6900
CCCTCGTGCG CTCTCCTGTT CCGACCCTGC CGCTTACCGG ATACCTGTCC GCCTTTCTCC    6960
CTTCGGGAAG CGTGGCGCTT TCTCAATGCT CACGCTGTAG GTATCTCAGT TCGGTGTAGG    7020
TCGTTCGCTC CAAGCTGGGC TGTGTGCACG AACCCCCCGT TCAGCCCGAC CGCTGCGCCT    7080
TATCCGGTAA CTATCGTCTT GAGTCCAACC CGGTAAGACA CGACTTATCG CCACTGGCAG    7140
CAGCCACTGG TAACAGGATT AGCAGAGCGA GGTATGTAGG CGGTGCTACA GAGTTCTTGA    7200
AGTGGTGGCC TAACTACGGC TACACTAGAA GGACAGTATT TGGTATCTGC GCTCTGCTGA    7260
AGCCAGTTAC CTTCGGAAAA AGAGTTGGTA GCTCTTGATC CGGCAAACAA ACCACCGCTG    7320
GTAGCGGTGG TTTTTTTGTT TGCAAGCAGC AGATTACGCG CAGAAAAAAA GGATCTCAAG    7380
AAGATCCTTT GATCTTTTCT ACGGGGTCTG ACGCTCAGTG GAACGAAAAC TCACGTTAAG    7440
GGATTTTGGT CATGAGATTA TCAAAAAGGA TCTTCACCTA GATCCTTTTA AATTAAAAAT    7500
GAAGTTTTAA ATCAATCTAA AGTATATATG AGTAAACTTG GTCTGACAGT TACCAATGCT    7560
TAATCAGTGA GGCACCTATC TCAGCGATCT GTCTATTTCG TTCATCCATA GTTGCCTGAC    7620
TCCCCGTCGT GTAGATAACT ACGATACGGG AGGGCTTACC ATCTGGCCCC AGTGCTGCAA    7680
```

```
TGATACCGCG AGACCCACGC TCACCGGCTC CAGATTTATC AGCAATAAAC CAGCCAGCCG      7740
GAAGGGCCGA GCGCAGAAGT GGTCCTGCAA CTTTATCCGC CTCCATCCAG TCTATTAATT      7800
GTTGCCGGGA AGCTAGAGTA AGTAGTTCGC CAGTTAATAG TTTGCGCAAC GTTGTTGCCA      7860
TTGCTACAGG CATCGTGGTG TCACGCTCGT CGTTTGGTAT GGCTTCATTC AGCTCCGGTT      7920
CCCAACGATC AAGGCGAGTT ACATGATCCC CCATGTTGTG CAAAAAAGCG GTTAGCTCCT      7980
TCGGTCCTCC GATCGTTGTC AGAAGTAAGT TGGCCGCAGT GTTATCACTC ATGGTTATGG      8040
CAGCACTGCA TAATTCTCTT ACTGTCATGC CATCCGTAAG ATGCTTTTCT GTGACTGGTG      8100
AGTACTCAAC CAAGTCATTC TGAGAATAGT GTATGCGGCG ACCGAGTTGC TCTTGCCCGG      8160
CGTCAATACG GGATAATACC GCGCCACATA GCAGAACTTT AAAAGTGCTC ATCATTGGAA      8220
AACGTTCTTC GGGGCGAAAA CTCTCAAGGA TCTTACCGCT GTTGAGATCC AGTTCGATGT      8280
AACCCACTCG TGCACCCAAC TGATCTTCAG CATCTTTTAC TTTCACCAGC GTTTCTGGGT      8340
GAGCAAAAAC AGGAAGGCAA AATGCCGCAA AAAAGGGAAT AAGGGCGACA CGGAAATGTT      8400
GAATACTCAT ACTCTTCCTT TTTCAATATT ATTGAAGCAT TTATCAGGGT TATTGTCTCA      8460
TGAGCGGATA CATATTTGAA TGTATTTAGA AAAATAAACA AATAGGGGTT CCGCGCACAT      8520
                                                                       8540
TTCCCCGAAA AGTGCCACCT
```

(2) INFORMATION FOR SEQ ID NO:3:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 9209 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: circular

    (ii) MOLECULE TYPE: DNA (genomic)

   (iii) HYPOTHETICAL: NO

    (iv) ANTI-SENSE: NO

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
GACGTCGCGG CCGCTCTAGG CCTCCAAAAA AGCCTCCTCA CTACTTCTGG AATAGCTCAG        60
AGGCCGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAAT TAGTCAGCCA TGCATGGGGC     - 120
GGAGAATGGG CGGAACTGGG CGGAGTTAGG GGCGGGATGG GCGGAGTTAG GGGCGGGACT       180
ATGGTTGCTG ACTAATTGAG ATGCATGCTT TGCATACTTC TGCCTGCTGG GGAGCCTGGG       240
GACTTTCCAC ACCTGGTTGC TGACTAATTG AGATGCATGC TTTGCATACT TCTGCCTGCT       300
GGGGAGCCTG GGGACTTTCC ACACCCTAAC TGACACACAT TCCACAGAAT TAATTCCCCT       360
AGTTATTAAT AGTAATCAAT TACGGGGTCA TTAGTTCATA GCCCATATAT GGAGTTCCGC       420
GTTACATAAC TTACGGTAAA TGGCCCGCCT GGCTGACCGC CCAACGACCC CGCCCATTG       480
ACGTCAATAA TGACGTATGT TCCCATAGTA ACGCCAATAG GGACTTTCCA TTGACGTCAA       540
TGGGTGGACT ATTTACGGTA AACTGCCCAC TTGGCAGTAC ATCAAGTGTA TCATATGCCA       600
```

- 71 -

```
AGTACGCCCC CTATTGACGT CAATGACGGT AAATGGCCCG CCTGGCATTA TGCCCAGTAC      660
ATGACCTTAT GGGACTTTCC TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC      720
ATGGTGATGC GGTTTTGGCA GTACATCAAT GGGCGTGGAT ACCGGTTTGA CTCACGCGGA      780
TTTCCAAGTC TCCACCCCAT TGACGTCAAT GGGAGTTTGT TTTGGCACCA AAATCAACGG      840
GACTTTCCAA AATGTCGTAA CAACTCCGCC CCATTGACGC AAATGGGCGG TAGGCGTGTA      900
CGGTGGGAGG TCTATATAAG CAGAGCTGGG TACGTGAACC GTCAGATCGC CTGGAGACGC      960
CATCACAGAT CTCTCACTAT GGATTTTCAG GTGCAGATTA TCAGCTTCCT GCTAATCAGT     1020
GCTTCAGTCA TAATGTCCAG AGGACAAATT GTTCTCTCCC AGTCTCCAGC AATCCTGTCT     1080
GCATCTCCAG GGGAGAAGGT CACAATGACT TGCAGGGCCA GCTCAAGTGT AAGTTACATC     1140
CACTGGTTCC AGCAGAAGCC AGGATCCTCC CCCAAACCCT GGATTTATGC CACATCCAAC     1200
CTGGCTTCTG GAGTCCCTGT TCGCTTCAGT GGCAGTGGGT CTGGGACTTC TTACTCTCTC     1260
ACAATCAGCA GAGTGGAGGC TGAAGATGCT GCCACTTATT ACTGCCAGCA GTGGACTAGT     1320
AACCCACCCA CGTTCGGAGG GGGGACCAAG CTGGAAATCA AACGTACGGT GGCTGCACCA     1380
TCTGTCTTCA TCTTCCCGCC ATCTGATGAG CAGTTGAAAT CTGGAACTGC CTCTGTTGTG     1440
TGCCTGCTGA ATAACTTCTA TCCCAGAGAG GCCAAAGTAC AGTGGAAGGT GGATAACGCC     1500
CTCCAATCGG GTAACTCCCA GGAGAGTGTC ACAGAGCAGG ACAGCAAGGA CAGCACCTAC     1560
AGCCTCAGCA GCACCCTGAC GCTGAGCAAA GCAGACTACG AGAAACACAA AGTCTACGCC     1620
TGCGAAGTCA CCCATCAGGG CCTGAGCTCG CCCGTCACAA AGAGCTTCAA CAGGGGAGAG     1680
TGTTGAATTC AGATCCGTTA ACGGTTACCA ACTACCTAGA CTGGATTCGT GACAACATGC     1740
GGCCGTGATA TCTACGTATG ATCAGCCTCG ACTGTGCCTT CTAGTTGCCA GCCATCTGTT     1800
GTTTGCCCCT CCCCCGTGCC TTCCTTGACC CTGGAAGGTG CCACTCCCAC TGTCCTTTCC     1860
TAATAAAATG AGGAAATTGC ATCGCATTGT CTGAGTAGGT GTCATTCTAT TCTGGGGGGT     1920
GGGGTGGGGC AGGACAGCAA GGGGGAGGAT TGGGAAGACA ATAGCAGGCA TGCTGGGGAT     1980
GCGGTGGGCT CTATGGAACC AGCTGGGGCT CGACAGCTAT GCCAAGTACG CCCCCTATTG     2040
ACGTCAATGA CGGTAAATGG CCCGCCTGGC ATTATGCCCA GTACATGACC TTATGGGACT     2100
TTCCTACTTG GCAGTACATC TACGTATTAG TCATCGCTAT TACCATGGTG ATGCGGTTTT     2160
GGCAGTACAT CAATGGGCGT GGATAGCGGT TTGACTCACG GGGATTTCCA AGTCTCCACC     2220
CCATTGACGT CAATGGGAGT TTGTTTTGGC ACCAAAATCA ACGGGACTTT CCAAAATGTC     2280
GTAACAACTC CGCCCCATTG ACGCAAATGG GCGGTAGGCG TGTACGGTGG GAGGTCTATA     2340
TAAGCAGAGC TGGGTACGTC CTCACATTCA GTGATCAGCA CTGAACACAG ACCCGTCGAC     2400
ATGGGTTGGA GCCTCATCTT GCTCTTCCTT GTCGCTGTTG CTACGCGTGT CCTGTCCCAG     2460
GTACAACTGC AGCAGCCTGG GGCTGAGCTG GTGAAGCCTG GGCCTCAGT GAAGATGTCC      2520
TGCAAGGCTT CTGGCTACAC ATTTACCAGT TACAATATGC ACTGGGTAAA ACAGACACCT     2580
GGTCGGGGCC TGGAATGGAT TGGAGCTATT TATCCCGGAA ATGGTGATAC TTCCTACAAT     2640
```

CAGAAGTTCA AAGGCAAGGC ●ATTGACT GCAGACAAAT CCTCCAGCAC     CTACATG    2700

CAGCTCAGCA GCCTGACATC TGAGGACTCT GCGGTCTATT ACTGTGCAAG ATCGACTTAC    2760

TACGGCGGTG ACTGGTACTT CAATGTCTGG GGCGCAGGGA CCACGGTCAC CGTCTCTGCA    2820

GCTAGCACCA AGGGCCCATC GGTCTTCCCC CTGGCACCCT CCTCCAAGAG CACCTCTGGG    2880

GGCACAGCGG CCCTGGGCTG CCTGGTCAAG GACTACTTCC CCGAACCGGT GACGGTGTCG    2940

TGGAACTCAG GCGCCCTGAC CAGCGGCGTG CACACCTTCC CGGCTGTCCT ACAGTCCTCA    3000

GGACTCTACT CCCTCAGCAG CGTGGTGACC GTGCCCTCCA GCAGCTTGGG CACCCAGACC    3060

TACATCTGCA ACGTGAATCA CAAGCCCAGC AACACCAAGG TGGACAAGAA AGCAGAGCCC    3120

AAATCTTGTG ACAAAACTCA CACATGCCCA CCGTGCCCAG CACCTGAACT CCTGGGGGGA    3180

CCGTCAGTCT TCCTCTTCCC CCCAAAACCC AAGGACACCC TCATGATCTC CCGGACCCCT    3240

GAGGTCACAT GCGTGGTGGT GGACGTGAGC CACGAAGACC CTGAGGTCAA GTTCAACTGG    3300

TACGTGGACG GCGTGGAGGT GCATAATGCC AAGACAAAGC CGCGGGAGGA GCAGTACAAC    3360

AGCACGTACC GTGTGGTCAG CGTCCTCACC GTCCTGCACC AGGACTGGCT GAATGGCAAG    3420

GAGTACAAGT GCAAGGTCTC CAACAAAGCC CTCCCAGCCC CCATCGAGAA AACCATCTCC    3480

AAAGCCAAAG GGCAGCCCCG AGAACCACAG GTGTACACCC TGCCCCCATC CCGGGATGAG    3540

CTGACCAAGA ACCAGGTCAG CCTGACCTGC CTGGTCAAAG GCTTCTATCC CAGCGACATC    3600

GCCGTGGAGT GGGAGAGCAA TGGGCAGCCG GAGAACAACT ACAAGACCAC GCCTCCCGTG    3660

CTGGACTCCG ACGGCTCCTT CTTCCTCTAC AGCAAGCTCA CCGTGGACAA GAGCAGGTGG    3720

CAGCAGGGGA ACGTCTTCTC ATGCTCCGTG ATGCATGAGG CTCTGCACAA CCACTACACG    3780

CAGAAGAGCC TCTCCCTGTC TCCGGGTAAA TGAGGATCCG TTAACGGTTA CCAACTACCT    3840

AGACTGGATT CGTGACAACA TGCGGCCGTG ATATCTACGT ATGATCAGCC TCGACTGTGC    3900

CTTCTAGTTG CCAGCCATCT GTTGTTTGCC CCTCCCCCGT GCCTTCCTTG ACCCTGGAAG    3960

GTGCCACTCC CACTGTCCTT TCCTAATAAA ATGAGGAAAT TGCATCGCAT TGTCTGAGTA    4020

GGTGTCATTC TATTCTGGGG GGTGGGGTGG GGCAGGACAG CAAGGGGGAG GATTGGGAAG    4080

ACAATAGCAG GCATGCTGGG GATGCGGTGG GCTCTATGGA ACCAGCTGGG GCTCGACAGC    4140

GCTGGATCTC CCGATCCCCA GCTTTGCTTC TCAATTTCTT ATTTGCATAA TGAGAAAAAA    4200

AGGAAAATTA ATTTTAACAC CAATTCAGTA GTTGATTGAG CAAATGCGTT GCCAAAAAGG    4260

ATGCTTTAGA GACAGTGTTC TCTGCACAGA TAAGGACAAA CATTATTCAG AGGGAGTACC    4320

CAGAGCTGAG ACTCCTAAGC CAGTGAGTGG CACAGCATTC TAGGGAGAAA TATGCTTGTC    4380

ATCACCGAAG CCTGATTCCG TAGAGCCACA CCTTGGTAAG GGCCAATCTG CTCACACAGG    4440

ATAGAGAGGG CAGGAGCCAG GGCAGAGCAT ATAAGGTGAG GTAGGATCAG TTGCTCCTCA    4500

CATTTGCTTC TGACATAGTT GTGTTGGGAG CTTGGATAGC TTGGACAGCT CAGGGCTGCG    4560

ATTTCGCGCC AAACTTGACG GCAATCCTAG CGTGAAGGCT GGTAGGATTT TATCCCCGCT    4620

GCCATCATGG TTCGACCATT GAACTGCATC GTCGCCGTGT CCCAAAATAT GGGGATTGGC    4680

```
AAGAACGGAG ACCTACCCTG ●TTCCGCTC AGGAACGAGT TCAAGTACTT ●AAGAATG      4740
ACCACAACCT CTTCAGTGGA AGGTAAACAG AATCTGGTGA TTATGGGTAG GAAAACCTGG     4800
TTCTCCATTC CTGAGAAGAA TCGACCTTTA AAGGACAGAA TTAATATAGT TCTCAGTAGA     4860
GAACTCAAAG AACCACCACG AGGAGCTCAT TTTCTTGCCA AAAGTTTGGA TGATGCCTTA     4920
AGACTTATTG AACAACCGGA ATTGGCAAGT AAAGTAGACA TGGTTTGGAT AGTCGGAGGC     4980
AGTTCTGTTT ACCAGGAAGC CATGAATCAA CCAGGCCACC TTAGACTCTT TGTGACAAGG     5040
ATCATGCAGG AATTTGAAAG TGACACGTTT TTCCCAGAAA TTGATTTGGG GAAATATAAA     5100
CTTCTCCCAG AATACCCAGG CGTCCTCTCT GAGGTCCAGG AGGAAAAAGG CATCAAGTAT     5160
AAGTTTGAAG TCTACGAGAA GAAAGACTAA CAGGAAGATG CTTTCAAGTT CTCTGCTCCC     5220
CTCCTAAAGC TATGCATTTT TATAAGACCA TGGGACTTTT GCTGGCTTTA GATCAGCCTC     5280
GACTGTGCCT TCTAGTTGCC AGCCATCTGT TGTTTGCCCC TCCCCCGTGC CTTCCTTGAC     5340
CCTGGAAGGT GCCACTCCCA CTGTCCTTTC CTAATAAAAT GAGGAAATTG CATCGCATTG     5400
TCTGAGTAGG TGTCATTCTA TTCTGGGGGG TGGGGTGGGG CAGGACAGCA AGGGGGAGGA     5460
TTGGGAAGAC AATAGCAGGC ATGCTGGGGA TGCGGTGGGC TCTATGGAAC CAGCTGGGGC     5520
TCGAGCTACT AGCTTTGCTT CTCAATTTCT TATTTGCATA ATGAGAAAAA AAGGAAAATT     5580
AATTTTAACA CCAATTCAGT AGTTGATTGA GCAAATGCGT TGCCAAAAAG GATGCTTTAG     5640
AGACAGTGTT CTCTGCACAG ATAAGGACAA ACATTATTCA GAGGGAGTAC CCAGAGCTGA     5700
GACTCCTAAG CCAGTGAGTG GCACAGCATT CTAGGGAGAA ATATGCTTGT CATCACCGAA     5760
GCCTGATTCC GTAGAGCCAC ACCTTGGTAA GGGCCAATCT GCTCACACAG GATAGAGAGG     5820
GCAGGAGCCA GGGCAGAGCA TATAAGGTGA GGTAGGATCA GTTGCTCCTC ACATTTGCTT     5880
CTGACATAGT TGTGTTGGGA GCTTGGATCG ATCCTCTATG GTTGAACAAG ATGGATTGCA     5940
CGCAGGTTCT CCGGCCGCTT GGGTGGAGAG GCTATTCGGC TATGACTGGG CACAACAGAC     6000
AATCGGCTGC TCTGATGCCG CCGTGTTCCG GCTGTCAGCG CAGGGGCGCC CGGTTCTTTT     6060
TGTCAAGACC GACCTGTCCG GTGCCCTGAA TGAACTGCAG GACGAGGCAG CGCGGCTATC     6120
GTGGCTGGCC ACGACGGGCG TTCCTTGCGC AGCTGTGCTC GACGTTGTCA CTGAAGCGGG     6180
AAGGGACTGG CTGCTATTGG GCGAAGTGCC GGGGCAGGAT CTCCTGTCAT CTCACCTTGC     6240
TCCTGCCGAG AAAGTATCCA TCATGGCTGA TGCAATGCGG CGGCTGCATA CGCTTGATCC     6300
GGCTACCTGC CCATTCGACC ACCAAGCGAA ACATCGCATC GAGCGAGCAC GTACTCGGAT     6360
GGAAGCCGGT CTTGTCGATC AGGATGATCT GGACGAAGAG CATCAGGGGC TCGCGCCAGC     6420
CGAACTGTTC GCCAGGCTCA AGGCGCGCAT GCCCGACGGC GAGGATCTCG TCGTGACCCA     6480
TGGCGATGCC TGCTTGCCGA ATATCATGGT GGAAAATGGC CGCTTTTCTG GATTCATCGA     6540
CTGTGGCCGG CTGGGTGTGG CGGACCGCTA TCAGGACATA GCGTTGGCTA CCCGTGATAT     6600
TGCTGAAGAG CTTGGCGGCG AATGGGCTGA CCGCTTCCTC GTGCTTTACG GTATCGCCGC     6660
TCCCGATTCG CAGCGCATCG CCTTCTATCG CCTTCTTGAC GAGTTCTTCT GAGCGGGACT     6720
```

- 74 -

CTGGGGTTCG AAATGACCGA ██AGCGACG CCCAACCTGC CATCACGAGA    CGATTCC    6780

ACCGCCGCCT TCTATGAAAG GTTGGGCTTC GGAATCGTTT TCCGGGACGC CGGCTGGATG    6840

ATCCTCCAGC GCGGGGATCT CATGCTGGAG TTCTTCGCCC ACCCCAACTT GTTTATTGCA    6900

GCTTATAATG GTTACAAATA AAGCAATAGC ATCACAAATT TCACAAATAA AGCATTTTTT    6960

TCACTGCATT CTAGTTGTGG TTTGTCCAAA CTCATCAATC TATCTTATCA TGTCTGGATC    7020

GCGGCCGCGA TCCCGTCGAG AGCTTGGCGT AATCATGGTC ATAGCTGTTT CCTGTGTGAA    7080

ATTGTTATCC GCTCACAATT CCACACAACA TACGAGCCGG AAGCATAAAG TGTAAAGCCT    7140

GGGGTGCCTA ATGAGTGAGC TAACTCACAT TAATTGCGTT GCGCTCACTG CCCGCTTTCC    7200

AGTCGGGAAA CCTGTCGTGC CAGCTGCATT AATGAATCGG CCAACGCGCG GGGAGAGGCG    7260

GTTTGCGTAT TGGGCGCTCT TCCGCTTCCT CGCTCACTGA CTCGCTGCGC TCGGTCGTTC    7320

GGCTGCGGCG AGCGGTATCA GCTCACTCAA AGGCGGTAAT ACGGTTATCC ACAGAATCAG    7380

GGGATAACGC AGGAAAGAAC ATGTGAGCAA AAGGCCAGCA AAAGGCCAGG AACCGTAAAA    7440

AGGCCGCGTT GCTGGCGTTT TTCCATAGGC TCCGCCCCCC TGACGAGCAT CACAAAAATC    7500

GACGCTCAAG TCAGAGGTGG CGAAACCCGA CAGGACTATA AAGATACCAG GCGTTTCCCC    7560

CTGGAAGCTC CCTCGTGCGC TCTCCTGTTC CGACCCTGCC GCTTACCGGA TACCTGTCCG    7620

CCTTTCTCCC TTCGGGAAGC GTGGCGCTTT CTCAATGCTC ACGCTGTAGG TATCTCAGTT    7680

CGGTGTAGGT CGTTCGCTCC AAGCTGGGCT GTGTGCACGA ACCCCCCGTT CAGCCCGACC    7740

GCTGCGCCTT ATCCGGTAAC TATCGTCTTG AGTCCAACCC GGTAAGACAC GACTTATCGC    7800

CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG    7860

AGTTCTTGAA GTGGTGGCCT AACTACGGCT ACACTAGAAG GACAGTATTT GGTATCTGCG    7920

CTCTGCTGAA GCCAGTTACC TTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAAACAAA    7980

CCACCGCTGG TAGCGGTGGT TTTTTTGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAAG    8040

GATCTCAAGA AGATCCTTTG ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAAACT    8100

CACGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTCACCTAG ATCCTTTTAA    8160

ATTAAAAATG AAGTTTTAAA TCAATCTAAA GTATATATGA GTAAACTTGG TCTGACAGTT    8220

ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTTCGT TCATCCATAG    8280

TTGCCTGACT CCCCGTCGTG TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA    8340

GTGCTGCAAT GATACCGCGA GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAACC    8400

AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG GTCCTGCAAC TTTATCCGCC TCCATCCAGT    8460

CTATTAATTG TTGCCGGGAA GCTAGAGTAA GTAGTTCGCC AGTTAATAGT TTGCGCAACG    8520

TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTTGGTATG GCTTCATTCA    8580

GCTCCGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG    8640

TTAGCTCCTT CGGTCCTCCG ATCGTTGTCA GAAGTAAGTT GGCCGCAGTG TTATCACTCA    8700

TGGTTATGGC AGCACTGCAT AATTCTCTTA CTGTCATGCC ATCCGTAAGA TGCTTTTCTG    8760

- 75 -

```
TGACTGGTGA GTACTCAACC ●TCATTCT GAGAATAGTG TATGCGGCGA ●AGTTGCT        8820

CTTGCCCGGC GTCAATACGG GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA        8880

TCATTGGAAA ACGTTCTTCG GGGCGAAAAC TCTCAAGGAT CTTACCGCTG TTGAGATCCA        8940

GTTCGATGTA ACCCACTCGT GCACCCAACT GATCTTCAGC ATCTTTTACT TTCACCAGCG        9000

TTTCTGGGTG AGCAAAAACA GGAAGGCAAA ATGCCGCAAA AAAGGGAATA AGGGCGACAC        9060

GGAAATGTTG AATACTCATA CTCTTCCTTT TTCAATATTA TTGAAGCATT TATCAGGGTT        9120

ATTGTCTCAT GAGCGGATAC ATATTTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTC        9180

                                                                        9209
CGCGCACATT TCCCCGAAAA GTGCCACCT
```

(2) INFORMATION FOR SEQ ID NO:4:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 47 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (iii) HYPOTHETICAL: NO

    (iv) ANTI-SENSE: NO

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
ATCACAGATC TCTCACCATG GATTTTCAGG TGCAGATTAT CAGCTTC          47
```

(2) INFORMATION FOR SEQ ID NO:5:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 30 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (iii) HYPOTHETICAL: NO

    (iv) ANTI-SENSE: YES

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
TGCAGCATCC GTACGTTTGA TTTCCAGCTT          30
```

(2) INFORMATION FOR SEQ ID NO:6:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 384 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: N

        (ix) FEATURE:
             (A) NAME/KEY: CDS
             (B) LOCATION: 1..384

        (ix) FEATURE:
             (A) NAME/KEY: mat_peptide
             (B) LOCATION: 67..384

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
ATG GAT TTT CAG GTG CAG ATT ATC AGC TTC CTG CTA ATC AGT GCT TCA        48
Met Asp Phe Gln Val Gln Ile Ile Ser Phe Leu Leu Ile Ser Ala Ser
-22     -20             -15             -10

GTC ATA ATG TCC AGA GGG CAA ATT GTT CTC TCC CAG TCT CCA GCA ATC        96
Val Ile Met Ser Arg Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile
 -5              -1  1               5                   10

CTG TCT GCA TCT CCA GGG GAG AAG GTC ACA ATG ACT TGC AGG GCC AGC       144
Leu Ser Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser
                15              20              25

TCA AGT GTA AGT TAC ATC CAC TGG TTC CAG CAG AAG CCA GGA TCC TCC       192
Ser Ser Val Ser Tyr Ile His Trp Phe Gln Gln Lys Pro Gly Ser Ser
                30              35              40

CCC AAA CCC TGG ATT TAT GCC ACA TCC AAC CTG GCT TCT GGA GTC CCT       240
Pro Lys Pro Trp Ile Tyr Ala Thr Ser Asn Leu Ala Ser Gly Val Pro
        45              50              55

GTT CGC TTC AGT GGC AGT GGG TCT GGG ACT TCT TAC TCT CTC ACA ATC       288
Val Arg Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile
    60              65              70

AGC AGA GTG GAG GCT GAA GAT GCT GCC ACT TAT TAC TGC CAG CAG TGG       336
Ser Arg Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp
75              80              85              90

ACT AGT AAC CCA CCC ACG TTC GGA GGG GGG ACC AAG CTG GAA ATC AAA       384
Thr Ser Asn Pro Pro Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys
            95              100             105
```

(2) INFORMATION FOR SEQ ID NO:7:

        (i) SEQUENCE CHARACTERISTICS:
             (A) LENGTH: 27 base pairs
             (B) TYPE: nucleic acid
             (C) STRANDEDNESS: single
             (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA (genomic)

        (iii) HYPOTHETICAL: NO

        (iv) ANTI-SENSE: NO

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GCGGCTCCCA CGCGTGTCCT GTCCCAG                                          27

(2) INFORMATION FOR SEQ ID NO:8:

        (i) SEQUENCE CHARACTERISTICS:

(A)  LENGTH: ⬤ ase pairs
        (B)  TYPE: nucleic acid
        (C)  STRANDEDNESS: single
        (D)  TOPOLOGY: linear

    (ii)  MOLECULE TYPE: DNA (genomic)

   (iii)  HYPOTHETICAL: NO

    (iv)  ANTI-SENSE: YES

    (ix)  FEATURE:
        (A)  NAME/KEY: misc_feature
        (B)  LOCATION: 3
        (D)  OTHER INFORMATION: /note= "Nucleotide 3 is N wherein N
is G or C."

    (ix)  FEATURE:
        (A)  NAME/KEY: misc_feature
        (B)  LOCATION: 18
        (D)  OTHER INFORMATION: /note= "Nucleotide 18 is N wherein
N is A or C."

    (ix)  FEATURE:
        (A)  NAME/KEY: misc_feature
        (B)  LOCATION: 19
        (D)  OTHER INFORMATION: /note= "Nucleotide 19 is N wherein
N is A or G."

    (ix)  FEATURE:
        (A)  NAME/KEY: misc_feature
        (B)  LOCATION: 25
        (D)  OTHER INFORMATION: /note= "Nucleotide 25 is N wherein
N is G or A."

    (xi)  SEQUENCE DESCRIPTION: SEQ ID NO:8:

GGNTGTTGTG CTAGCTGNNG AGACNGTGA                                    29


(2)  INFORMATION FOR SEQ ID NO:9:

     (i)  SEQUENCE CHARACTERISTICS:
        (A)  LENGTH: 420 base pairs
        (B)  TYPE: nucleic acid
        (C)  STRANDEDNESS: single
        (D)  TOPOLOGY: linear

    (ii)  MOLECULE TYPE: DNA (genomic)

   (iii)  HYPOTHETICAL: NO

    (iv)  ANTI-SENSE: NO

    (ix)  FEATURE:
        (A)  NAME/KEY: CDS
        (B)  LOCATION: 1..420

    (ix)  FEATURE:
        (A)  NAME/KEY: mat_peptide
        (B)  LOCATION: 58..420

- 78 -

```
ATG GGT TGG AGC CTC ATC TTG CTC TTC CTT GTC GCT GTT GCT ACG CGT        48
Met Gly Trp Ser Leu Ile Leu Leu Phe Leu Val Ala Val Ala Thr Arg
-19              -15                   -10                   -5

GTC CTG TCC CAG GTA CAA CTG CAG CAG CCT GGG GCT GAG CTG GTG AAG        96
Val Leu Ser Gln Val Gln Leu Gln Gln Pro Gly Ala Glu Leu Val Lys
        -1   1               5                      10

CCT GGG GCC TCA GTG AAG ATG TCC TGC AAG GCT TCT GGC TAC ACA TTT       144
Pro Gly Ala Ser Val Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe
        15              20                  25

ACC AGT TAC AAT ATG CAC TGG GTA AAA CAG ACA CCT GGT CGG GGC CTG       192
Thr Ser Tyr Asn Met His Trp Val Lys Gln Thr Pro Gly Arg Gly Leu
    30              35                  40                  45

GAA TGG ATT GGA GCT ATT TAT CCC GGA AAT GGT GAT ACT TCC TAC AAT       240
Glu Trp Ile Gly Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn
            50                  55                  60

CAG AAG TTC AAA GGC AAG GCC ACA TTG ACT GCA GAC AAA TCC TCC AGC       288
Gln Lys Phe Lys Gly Lys Ala Thr Leu Thr Ala Asp Lys Ser Ser Ser
            65                  70                  75

ACA GCC TAC ATG CAG CTC AGC AGC CTG ACA TCT GAG GAC TCT GCG GTC       336
Thr Ala Tyr Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val
            80                  85                  90

TAT TAC TGT GCA AGA TCG ACT TAC TAC GGC GGT GAC TGG TAC TTC AAT       384
Tyr Tyr Cys Ala Arg Ser Thr Tyr Tyr Gly Gly Asp Trp Tyr Phe Asn
        95                  100                 105

GTC TGG GGC GCA GGG ACC ACG GTC ACC GTC TCT GCA                       420
Val Trp Gly Ala Gly Thr Thr Val Thr Val Ser Ala
110                 115                 120
```